

50325-0114 (1115 / 33981)

Patent

UNITED STATES PATENT APPLICATION

FOR

SELECTIVELY PASSING NETWORK ADDRESSES THROUGH A SERVER

INVENTORS:

DAVID N. STILL
MINQI DENG
THOMAS A. RODEN

PREPARED BY:

HICKMAN PALERMO TRUONG & BECKER LLP
1600 WILLOW STREET
SAN JOSE, CA 95125-5106
(408) 414-1080

EXPRESS MAIL CERTIFICATE OF MAILING

"Express Mail" mailing label number EL624353556US

Date of Deposit August 9, 2000

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

Casey Moore

(Typed or printed name of person mailing paper or fee)

Casey Moore

(Signature of person mailing paper or fee)

00636392-000000

SELECTIVELY PASSING NETWORK ADDRESSES THROUGH A SERVER

RELATED APPLICATION

This application is a Continuation-in-part of co-pending application Ser. No. 09/225,909, filed January 05, 1999, "Selectively Forced Re-direction of Network Traffic," by D. Still et al., and domestic priority is hereby claimed from such application.

FIELD OF THE INVENTION

The present invention generally relates to data processing. The invention relates more specifically to selectively passing network addresses through a server in a networked computer system.

BACKGROUND OF THE INVENTION

In a client/server computing environment in which clients communicate with servers over non-secure networks, enabling only authorized clients to receive services from the servers presents a variety of technical problems. In one approach, identifying information is collected from each client, and only known clients are given services from a server. In one implementation of this approach, a host that originates a request for services must provide its source IP address to the receiving server.

However, in this approach, there is a need to minimize the number of times that the client identifying information travels across the network, in order to prevent interception and spoofing. For example, allowing the IP address to be passed explicitly as part of a secure protocol message could represent a security weakness; a client could intentionally supply an incorrect address to gain greater access. Unfortunately, not allowing the address to be passed explicitly at all means that important identifying information is not available to the receiving server.

FIG. 1A is a simplified block diagram of a client/server system in which this general problem may arise. One or more clients 1112A, 1112B are coupled to a network 1104. Each

006080" 26E9E960

client 1112A, 1112B is a network end station device such as a personal computer, workstation, or the equivalent. Alternatively, each client 1112A, 1112B is a process, such as a standard Web browser. The term "originating host" is used herein to refer to any of clients 1112A, 1112B. Although only two clients are shown in FIG. 1A, in a practical system there
5 may be any number of such clients.

Network 1104 is one or more local area networks, wide area networks, or internetworks, alone or in combination. In a preferred embodiment, network 1104 is the global, packet-switched internetwork known as the Internet.

An HTTP server 1130 is coupled to network 1104 in a position logically separate
10 from clients 1112A, 1112B. HTTP server 1130 comprises one or more servers or software elements that can respond to client requests received in the Hypertext Transfer Protocol (HTTP). HTTP server 1130 may have one or more associated application servers that provide other services.

HTTP server 1130 stores or manages session data 1140, which comprises a plurality
15 of records that identify clients that are authorized to receive services from the HTTP server. In an embodiment, session data 1140 is created and stored by HTTP server 1130 for the purpose of uniquely identifying clients that are authorized to access HTTP server 1130, its resources, or other associated servers and resources. Each record in session data 1140 includes a key 1141 that includes, among other data, client IP address values 1144A, 1144B
20 associated or paired with random number values 1142A, 1142B. For purposes of illustrating an example, two (2) keys 1141 are shown in FIG. 1A, however, in a practical system there may be any number of keys or other information in session data 1140.

Each IP address value 1144A, 1144B normally represents an Internet Protocol (IP) address that is pre-assigned to and uniquely associated with one of the clients 1112A, 1112B.
25 In conventional usage, IP addresses are uniquely associated with specific client hardware such as a particular workstation or personal computer. However, such a hardware device may

00000000" 26E9E9E0

execute a plurality of instances of client applications, such as Web browsers. Accordingly, random values 1142A, 1142B may be used in order to uniquely identify more than one instance of a browser that is running on the same physical machine. Random values 1142A, 1142B are generated by HTTP server 1130 when a client having an authorized IP address
5 connects to the HTTP server.

FIG. 1B is a simplified block diagram of a system similar to that of FIG. 1A, in which like numbered blocks represent like elements. In FIG. 1B, however, a proxy server 1108 is coupled to network 1104 for the purpose of providing services that complement, but are not available in, an application server 1106, which is also coupled to network 1104.

10 Proxy servers are useful for merging functionality from different servers. For example, assume that application server 1106 offers Secure Sockets Layer functions, but not servlet capabilities, and proxy server 1108 can provide servlet functions. In this case, proxy server 1108 can act as proxy for application server 1106. Proxy server 1108 receives such requests from clients 1112A, 1112B and can respond to them. However, to carry out a
15 response, proxy server 1108 may need a service or information from application server 1106, by communication over logical path 1110. When application server 1106 receives a service request from proxy server 1108, the application server stores the IP address of the proxy server as part of a key of a record in session data 1140.

20 In a security scheme that requires the source IP address of the originating host, a system that includes a proxy server can result in problems. In particular, security can be compromised because the receiving server (e.g., application server 1106) always receives the IP address of the proxy server 1108 with a service request, rather than the IP address of the clients 1112A, 1112B that originate service requests.

25 One responsive measure is not allowing the address to be passed explicitly at all in the protocol, however, in that case important identifying information is not available to the

receiving server. If the address is passed, security is reduced because all accesses appear to originate from the proxy server.

Based on the foregoing, there is a need in this field for a way to pass originating client or host network address information through a proxy server to a receiving server, in a secure
5 manner.

There is a specific need for a way to pass the IP address of a Web client through a proxy server to an HTTP server or application server, in a secure protocol that does not always allow the IP address to be passed.

In the client/server computing environment a need may arise to enable a client to
10 communicate with two servers, each of which provide functions that represent a portion of a service desired by the client. In some cases, it is desirable to permit the client to communicate with only a first one of the two servers and to prevent direct contact between the client and the other of the two servers.

FIG. 1C is a block diagram of a networked computer system in which the foregoing
15 general problem may arise. Client 102 is a computer device such as a workstation, server, router, or switch. Client 102 is coupled to network 104, which is an interconnected combination of computers, terminals, or peripherals used to provide communications between two or more points. A first server 106 and a second server 108 are coupled to network 104, logically separated from client 102. In one embodiment, client 102, network
20 104, and servers 106, 108 communicate using TCP/IP network protocols, and using HTTP protocol messages that comprise requests and responses. Such protocols are exemplary and not required.

Client 102 may communicate an HTTP request for a service to network 104. The request includes a name or other identifier of server 106, which client 102 expects to provide
25 the requested service. Network 104 locates server 106 and routes the request to server 106, thereby establishing a logical connection 110 from client 102 to the server. Server 106

determines that it cannot provide the function or service solicited in the request, but that server 108 can provide the function or service. Accordingly, software elements in server 106 automatically divert or "redirect" the request to server 108 over a logical connection 112, which may physically travel through the network 104. Server 108 processes the request, generates a response message, and sends the response message back to client 102 over logical connection 114, which may pass through network 104.

In this scenario, when the client and the servers use HTTP, the response message may include a document formatted using a structured markup language, such as HTML. The HTML document may contain hyperlinks or other references to resources within server 108, or other servers or network elements. As a result, client 102 may select one of the hyperlinks or references, and thereby attempt to request a service of server 108 directly along connection 114. Server 106 would not be involved in processing such a request.

This result is undesirable in several circumstances. The server 106 may have been designated as authoritative for certain kinds of transactions. The server 106 may have redirected the original request to server 108 solely because server 106 cannot directly process the request, whereas server 108 can, but server 106 may need to remain in control of the overall transaction. For example, server 106 may have redirected the original request to server 108 just to carry out a specialized or subordinate task, although server 106 remains responsible for the total transaction or for presenting a consistent interface to the client 102.

Accordingly, there is a need in this field for a mechanism that allows a client request to be redirected from a first server to a second server, while keeping the first server in control of subsequent requests by the client for services of the second server.

In particular, there is a need for a mechanism that prevents the client from directly communicating with the second server even after the first server has redirected a request of the client to the second server.

SUMMARY OF THE INVENTION

The foregoing needs and objects, and other needs and objects that will become apparent from the following description, are achieved by the present invention, which comprises, in one aspect, a method of securely communicating a network address of a client

5 that issues service requests to a first server that proxies the service requests for a second server. A network address of the client is received. A processor determines whether a first network address of the first server is equal to a second network address of the second server. The network address of the client is sent from the first server to the second server in a secure request message only when the first network address of the first server is equal to the second

10 network address of the second server. Accordingly, a secure communications protocol is provided in which an address of a requesting client, e.g., an IP address, is passed in the protocol only among a responding server and its proxy, thereby preventing interception of the client IP address by unauthorized processes. By enforcing a policy that permits the network address of an originating host to pass from a first server to a second server only when the

15 network address of the second server meets specified criteria (e.g., it is the same network address as that of the first server), the originating host address can be passed securely through a proxy server.

006080-2629296

[illegible]

5 FIG. 1A is a simplified block diagram of a networked computer system;
 FIG. 1B is a simplified block diagram of a networked computer system that includes a
 proxy server;

FIG. 2 is a block diagram of a networked computer system according to an

FIG. 3A is a flow diagram of a process of selectively forced redirection of network

FIG. 3C is a flow diagram of a method of selectively allowing a source network

FIG. 4 is a block diagram of a computer system with which an embodiment of the invention may be implemented.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

A method and apparatus for selectively passing network addresses through a server is described. In the following description, for the purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the present invention. It
5 will be apparent, however, to one skilled in the art that the present invention may be practiced without these specific details. In other instances, well-known structures and devices are shown in block diagram form in order to avoid unnecessarily obscuring the present invention.

-- SELECTIVELY PASSING NETWORK ADDRESSES

FIG. 3C is a flow diagram of a method of selectively allowing a source network
10 address to pass through a server.

In block 340, a network address is received from a host that is originating a service request. For example, application server 1106 may receive a Universal Resource Locator (URL) or HTML POST form that includes a service request message, in which one parameter or field value is the IP address of the requesting host. The requesting host may be one of the
15 clients 1112A, 1112B, a Web browser executed by one of them, etc.

In block 342, a determination is made whether the network address of a destination server specified in the request is the same as the network address of the current server. In one embodiment, the service request message is received at a proxy server that proxies requests for the application server. Computer program code or other software or hardware elements
20 determine whether an IP address of a server specified in the request message is identical to the IP address of the proxy server. This will occur, for example, when the proxy server and application server are hosted on the same computer hardware. If the service request of the host is provided in the form of a URL, then block 342 may include carrying out ancillary steps, e.g., performing a DNS lookup of the IP address that corresponds to the hostname
25 portion of the URL, etc.

In block 344, if the test of block 342 yields a positive result, then the network address of the originating host is passed in a service request message from the current server to the destination server. Block 344 may involve creating an HTTP request message or URL that includes the client IP address. The service request message is forwarded or otherwise
5 communicated to the destination server in block 346.

Sub AI
If the test of block 342 yields a negative result, then the originating host is not allowed to send its network address to the destination server. In one embodiment, an error response message is sent back to the originating host, as shown by block 358. Block 358 may involve returning an HTTP response message with an error code, or any other suitable means
10 for signaling from the proxy server to the originating host that an error has occurred, or that security rules prohibit forwarding or responding to the request.

Accordingly, by enforcing a policy that permits the network address of an originating host to pass from a first server to a second server only when the network address of the second server meets specified criteria (e.g., it is the same network address as that of the first
15 server), the originating host address can be passed securely through a proxy server. As long as closed criteria are used for specifying when explicit network address passing is allowed, the address passing system can be secure.

-- SELECTIVELY FORCED REDIRECTION OF NETWORK TRAFFIC

FIG. 2 is a block diagram of a networked computer system in which a method of
20 selectively forced redirection of network traffic may be carried out.

As in the system of FIG. 1, a client 102 communicates with a network 104. In the preferred embodiment, client 102 executes a Web browser 103 and communicates with network 104 using HTTP requests and responses. A first server 208 and a second server 206 are coupled to network 104 and provide one or more network resources that can be accessed
25 by the browser 103 using appropriate HTTP requests. For clarity, FIG. 2 does not show all possible logical data communication paths that may be established between client 102 and

servers 206, 208. Path 209 is a logical communication path that may be used when server 206 and server 208 communicate with one another.

Server 206 is coupled to a Server Proxy 210, which may be one or more software or hardware elements or processes contained within or communicating with server 206. In one
5 embodiment, Server Proxy 210 is coupled to and interacts with a stream tokenizer 212 and an attribute list 214, however, these elements are not required.

Server 208 is coupled to or includes at least one resource 216 that is used by client 102. The resource 216 may be, for example, a Web page or a Web application, such as an interactive electronic commerce application. Similarly, server 206 is coupled to a resource
10 218, which may be a Web application or similar program.

In this configuration, Server Proxy 210 is configured to interact with server 206 to selectively force redirection of network traffic directed to server 208 from that server to server 206. In one embodiment, Server Proxy 210 is one or more software elements that are installed on and executed by server 206. In a preferred embodiment, Server Proxy 210 is
15 configured with computer program instructions that cause the system of FIG. 2 to operate generally as follows.

Client 102 submits a request for resource 218 to the network 104. The request may be submitted in the form of a URL that includes the name of server 206, which hosts resource 218. The term "host" means to store or execute a resource, or manage execution of a resource
20 located elsewhere.

Through means that are well known in this field, such as by using the DNS system, network 104 determines the location of server 206 and routes the client's request to that server. Server 206 receives and processes all requests by always redirecting them to server 208 over logical path 209. The redirection may be done selectively for particular types of
25 requests or may be done for all requests received by server 206.

A CGI script that invokes Server Proxy 210 is created and installed on the server that acts as proxy, for example, server 206 of FIG. 2. The CGI script is created and installed according to the requirements of that Web server. In addition to the regular CGI environment variables passed by the script, an 'actualServer' variable is passed, to tell Server Proxy 210 which server it is acting as a proxy for. A sample script is set forth in Table 1.

TABLE 1 -- EXAMPLE SCRIPT THAT INVOKES SERVER PROXY

```

10  #!/bin/sh
    BASE=/export/home/k2
    JBASE=/export/home/k2/java
    JAVA_HOME=$BASE/java
    JRE_TOP=$JAVA_HOME

15  JRE_CLASSES=$JRE_TOP/lib/rt.jar:$JRE_TOP/lib/i18n.jar:$JRE_TOP/lib/class
    es.jar:$JRE_TOP/lib/classes.zip:$JRE_TOP/classes

    CLASSPATH=/export/home/dstill/classes:$CLASSPATH:/export/home/k2/DBSer
    ver/DBServer.zip:/export/home/k2/FastAdmin/Acme.zip:/export/home/k2/FastAd
    min/fastadmin.zip:$WEBLOGIC/classes:$JRE_CLASSES
20  PATH=$PATH:$JAVA_HOME/bin
    export JAVA_HOME
    export JRE_TOP
    export JRE_CLASSES
    export CLASSPATH
25  export PATH
    cd $BASE/ns-home/docs/CScgi
    $JBASE/bin/java \
    -Dcgi.content_type=$CONTENT_TYPE \
30  -Dcgi.content_length=$CONTENT_LENGTH \
    -Dcgi.request_method=$REQUEST_METHOD \
    -Dcgi.query_string=$QUERY_STRING \
    -Dcgi.server_name=$SERVER_NAME \
    -Dcgi.server_port=$SERVER_PORT \
35  -Dcgi.script_name=$SCRIPT_NAME \
    -Dcgi.path_info=$PATH_INFO \
    -DactualServer="aladdin:9095" \
    cisco.ciscosecure.fastadmin.cgi.ServerProxy

```

-- STREAM TOKENIZER

Stream tokenizer 212 is a program or process that receives a text file as input, such as an HTML file, and produces a stream of tokens that represent tags and attributes within the contents of the input file. The tokens are produced according to token selection criteria that are coded into the stream tokenizer 212. An example of a suitable program that can be used for stream tokenizer 212 is the method `java.io.StreamTokenizer`, which is part of the Java Development Kit version 1.1.5, commercially available from Sun Microsystems, Inc. In off-the-shelf form, this Stream Tokenizer method is hard coded to treat character “\” in strings as an escape character. In HTML, however, the character “\” is not defined as an escape character and needs to be part of appropriate tokens. Accordingly, in the preferred embodiment, the method `java.io.StreamTokenizer` of the JDK is modified to enable such characters to appear in tokens. This may be done by commenting out the code that handles escape functions.

-- METHOD OF SELECTIVELY FORCING REDIRECTION OF NETWORK

TRAFFIC

FIG. 3A is a flow diagram of one embodiment of a method of selectively forcing redirection of network traffic. For convenience, the method of FIG. 3A is described below in connection with the structural elements of FIG. 2, however, the method of FIG. 3A does not require use of such a structure.

In block 302, a client request is received. For example, client 102 submits a request for resource 218 to the network 104 in the form of a URL that includes the name of server 206. Network 104 determines the location of server 206 and routes the client’s request to that server. In block 304, the client request is automatically redirected to another server that actually processes the request and generates a response. For example, server 206 redirects the client request it to server 208 over logical path 209.

In block 322, the process examines each token to identify each relative URL, or each explicit reference to a server. In block 324, when a token contains a relative URL or explicit server name reference, each attribute of the token is matched to a list of attributes that identify the server that created the results. The list of attributes may be a list of domain names, IP addresses, or other identifiers. Block 322 and block 324 may be carried out by
5 Server Proxy 210 in coordination with stream tokenizer 212 and attribute list 214. For example, attribute list 214 includes: HREF; SRC; ACTION. In block 326, the process tests whether a match is found. If there is a match, then the attribute of the current token is considered a candidate for translation into a modified reference.

10 For example, when a token is associated with a URL that is expressed in the HTML relative URL form, meaning that it refers to a directory of the server that created the results, then that token is identified as a candidate for translation. Also, if a token contains an express reference to that server, such as a URL that literally contains the server's name, then that token is also identified as a candidate for translation.

15 Tokens that are not candidates for translation are written or otherwise output into the output result file, as shown in block 330.

When a token is identified as a candidate for translation, then control is passed to block 328, in which a reference to a script is prepended to the URL attribute of the token. For example, block 328 may involve prepending the name of a CGI script, which implements the
20 process of FIG. 3A and FIG. 3B, to the URL attribute of the token. Control is then passed to block 330 in which the token is output.

Subsequently, when a user of client 102 clicks on a hyperlink or other reference of the modified result, the hyperlink will always direct browser 103 to server 206 or to Server Proxy 210. Server 206 will then redirect the browser 103 to server 208, which processes the
25 subsequent request. If HTML is generated again, then Server Proxy 210 will translate any references to server 208 into references to server 206. In this way, server 206 retains control

over the overall transaction with browser 103 and prevents the browser or client from interacting directly with server 208.

-- HARDWARE OVERVIEW

FIG. 4 is a block diagram that illustrates a computer system 400 upon which an embodiment of the invention may be implemented. Computer system 400 includes a bus 402 or other communication mechanism for communicating information, and a processor 404 coupled with bus 402 for processing information. Computer system 400 also includes a main memory 406, such as a random access memory (RAM) or other dynamic storage device, coupled to bus 402 for storing information and instructions to be executed by processor 404.

Main memory 406 also may be used for storing temporary variables or other intermediate information during execution of instructions to be executed by processor 404. Computer system 400 further includes a read only memory (ROM) 408 or other static storage device coupled to bus 402 for storing static information and instructions for processor 404. A storage device 410, such as a magnetic disk or optical disk, is provided and coupled to bus 402 for storing information and instructions.

Computer system 400 may be coupled via bus 402 to a display 412, such as a cathode ray tube (CRT), for displaying information to a computer user. An input device 414, including alphanumeric and other keys, is coupled to bus 402 for communicating information and command selections to processor 404. Another type of user input device is cursor control 416, such as a mouse, a trackball, or cursor direction keys for communicating direction information and command selections to processor 404 and for controlling cursor movement on display 412. This input device typically has two degrees of freedom in two axes, a first axis (e.g., x) and a second axis (e.g., y), that allows the device to specify positions in a plane.

The invention is related to the use of computer system 400 for selectively forced redirection of network traffic. According to one embodiment of the invention, selectively

006030" 25E9E960

forced redirection of network traffic is provided by computer system 400 in response to processor 404 executing one or more sequences of one or more instructions contained in main memory 406. Such instructions may be read into main memory 406 from another computer-readable medium, such as storage device 410. Execution of the sequences of instructions contained in main memory 406 causes processor 404 to perform the process steps described herein. In alternative embodiments, hard-wired circuitry may be used in place of or in combination with software instructions to implement the invention. Thus, embodiments of the invention are not limited to any specific combination of hardware circuitry and software.

The term "computer-readable medium" as used herein refers to any medium that participates in providing instructions to processor 404 for execution. Such a medium may take many forms, including but not limited to, non-volatile media, volatile media, and transmission media. Non-volatile media includes, for example, optical or magnetic disks, such as storage device 410. Volatile media includes dynamic memory, such as main memory 406. Transmission media includes coaxial cables, copper wire and fiber optics, including the wires that comprise bus 402. Transmission media can also take the form of acoustic or light waves, such as those generated during radio-wave and infra-red data communications.

Common forms of computer-readable media include, for example, a floppy disk, a flexible disk, hard disk, magnetic tape, or any other magnetic medium, a CD-ROM, any other optical medium, punchcards, papertape, any other physical medium with patterns of holes, a RAM, a PROM, and EPROM, a FLASH-EPROM, any other memory chip or cartridge, a carrier wave as described hereinafter, or any other medium from which a computer can read.

Various forms of computer readable media may be involved in carrying one or more sequences of one or more instructions to processor 404 for execution. For example, the instructions may initially be carried on a magnetic disk of a remote computer. The remote computer can load the instructions into its dynamic memory and send the instructions over a telephone line using a modem. A modem local to computer system 400 can receive the data

on the telephone line and use an infrared transmitter to convert the data to an infrared signal. An infrared detector can receive the data carried in the infrared signal and appropriate circuitry can place the data on bus 402. Bus 402 carries the data to main memory 406, from which processor 404 retrieves and executes the instructions. The instructions received by
5 main memory 406 may optionally be stored on storage device 410 either before or after execution by processor 404.

Computer system 400 also includes a communication interface 418 coupled to bus 402. Communication interface 418 provides a two-way data communication coupling to a network link 420 that is connected to a local network 422. For example, communication
10 interface 418 may be an integrated services digital network (ISDN) card or a modem to provide a data communication connection to a corresponding type of telephone line. As another example, communication interface 418 may be a local area network (LAN) card to provide a data communication connection to a compatible LAN. Wireless links may also be implemented. In any such implementation, communication interface 418 sends and receives
15 electrical, electromagnetic or optical signals that carry digital data streams representing various types of information.

Network link 420 typically provides data communication through one or more networks to other data devices. For example, network link 420 may provide a connection through local network 422 to a host computer 424 or to data equipment operated by an
20 Internet Service Provider (ISP) 426. ISP 426 in turn provides data communication services through the worldwide packet data communication network now commonly referred to as the "Internet" 428. Local network 422 and Internet 428 both use electrical, electromagnetic or optical signals that carry digital data streams. The signals through the various networks and the signals on network link 420 and through communication interface 418, which carry the
25 digital data to and from computer system 400, are exemplary forms of carrier waves transporting the information.

Computer system 400 can send messages and receive data, including program code, through the network(s), network link 420 and communication interface 418. In the Internet example, a server 430 might transmit a requested code for an application program through Internet 428, ISP 426, local network 422 and communication interface 418. In accordance
5 with the invention, one such downloaded application provides for selectively forced redirection of network traffic as described herein.

The received code may be executed by processor 404 as it is received, and/or stored in storage device 410, or other non-volatile storage for later execution. In this manner, computer system 400 may obtain application code in the form of a carrier wave.

10 -- SCOPE

In the foregoing specification, the invention has been described with reference to specific embodiments thereof. It will, however, be evident that various modifications and changes may be made thereto without departing from the broader spirit and scope of the invention. For example, the invention and claims do not require use of HTML, HTTP, or
15 TCP/IP. These are merely convenient examples of formats and protocols that may be used. The invention encompasses and is fully applicable to, for example, results that are expressed in any hypertext format, or any other format that may contain references to the thing that created it.

The specification and drawings are, accordingly, to be regarded in an illustrative
20 rather than a restrictive sense.
